PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Elucidating symptoms of COVID-19 illness in the Arizona CoVHORT, a longitudinal cohort study
AUTHORS	Khan, Sana; Farland, Leslie V.; Catalfamo, Collin; Austhof, Erika; Bell, Melanie L.; Chen, Zhao; Cordova-Marks, Felina; Ernst, Kacey; Garcia-Filion, Pamela; Heslin, Kelly M.; Hoskinson, Joshua; Jehn, Megan; Joseph, Emily C.S.; Kelley, Connor; Klimentidis, Yann; Russo Carroll, Stephanie; Kohler, Lindsay; Pogreba-Brown, Kristen; Jacobs, Elizabeth

VERSION 1 – REVIEW

REVIEWER	Seidelmann, Sara Columbia University
REVIEW RETURNED	10-Aug-2021

GENERAL COMMENTS	Major comments: 1) The authors state in the abstract that their main objective is to "aid in earlier identification of SARS-CoV-2 infection" by elucidating a symptom profile for SARS-CoV-2 infection. Yet they haven't performed basic analyses to achieve their objective such as sensitivity or specificity analyses. 2) Relevance. The authors cite a few similar studies to their work but state that "However, a major limitation of studies conducted to date is the lack of comparison of patient-reported symptoms to those of uninfected individuals." The authors attention should be drawn to two Cochrane reviews which have reviewed exactly that. The most recent Cochrane review, published in Feb 2021, entitled, "Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19" included 44 studies including 26,884 participants in total. The authors should justify how this study could possible add to the vast literature that already exists on this topic. Additionally, the prior data has suggested that specific symptoms are not a very good predictors of positive test. How could this relatively small data set change that conclusion and should it? 3) The text states: "Participants were asked if they had tested for the virus that causes COVID-19 with a nasal swab, throat swab, or saliva test since January 2020. Participants were classified as untested, positive or negative based on their results. Therefore, the Covid cases in this study are not lab confirmed, they are self-reported. All parts of the manuscript that refer to "lab-confirmed Covid" should be replaced with "self-reported positive Covid test."

	4) The “untested participants” should be split by those that answered “yes” to the question “Since January 1, have you experienced a sudden illness that led you to believe that you had Covid 19” and those that did not in all analyses and tables. 5) Sensitivity and Specificity analyses should be performed Minor: 6) Please report p-values for BMI and Covid symptoms analyzed as continuous variables. 7) Please add BMI and smoking to the multivariable model 8) The text states: Participants, regardless of COVID-19 test status, were asked, “Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?” If they answered “yes”, all participants, regardless of case status, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. “If they answered “yes” should be omitted and the Sentence should start, “All participants, regardless of case status, were asked… 9) Limitations should be expanded to include the self-reported nature of this data. Question order bias should be addressed (ie were those that answered "no" to Covid illness or to a test going to be as liking to describe symptoms), as should non-response bias.

REVIEWER	Kirchner, Marietta UniversitätsKlinikum Heidelberg, Medical Biometry
REVIEW RETURNED	15-Sep-2021

GENERAL COMMENTS	General aspects If I understood it correctly, the cohort was launched at the end of May 2020 and participants were asked about the time period since January 1, so participants had to remember which symptoms they had experienced and when. One can imagine that participants with confirmed diagnosis some time ago does not remember as good as participants with a confirmed diagnosis shortly before being included in the study. Have you asked for the date of the test and was this included in the analysis? Is this the variable „days since symtoms began“ (Table 2, variable not defined in the text)? Additionally, subjects could have had two tests since January 1 (e.g. two negative ones). Have you asked for that? It seems to me that not all participants were asked about symptoms but only those who believe to have had COVID19 as written on page 6 (lines 5/6). If this is the case then the objective has to be restriced to this subgroup as one only knows about symptoms from that group. One has to show the frequency for answering yes to that question stratified by case status. Additionally, I was wondering if participants with no symptoms from Table 2 all belong to those answering ‚no‘ but never indicated directly to have had no symptoms. And for the untested group: I think that nearlly all here answered ‚no‘ and were not aksed about

	symptoms as one would have had performed a test if one believe to have COVID19. Please specify. Abstract - Restructure the sections in the abstract as the results should be described solely in the Results section and not in the Objective, Participants, Outcome Measure section. For example, the Outcome Measure section should describe the outcomes (the clear definition of the primary and secondary outcomes is also missing in the main text; this has to be added). Statistical analysis section - Lines 18/19: one has to specify that differences in reported symptoms are identified and that this is only between positive and untested as well as positive and negative. The comparison of untested and negative is not reported. - Lines 22/23: it has to be specified that the comparison by ordered logistic regression is only with repect to the COVID-19 positive study participants - I do not see that nonparametric analogs were used. Please indicate here when and with respect to which analysis would you use which nonparametric analog. - Lines 22/23: what was the reason for these symptom categories (0, 1-6, 7-9, 10-16)? - Line 26: one has to add that age, sex and ethnicity were included additionally in the logistic regression model. Please justify why those variables were considered as confounder and not e.g. BMI. - Indicate how missing data were handeled - Indicate that this is an explorative analysis and that p-values have to be interpreted descriptively (no confirmatory value). Why reporting p-values for the ordered logistic regression and not for the logistic regression. Indicate that Odds Ratios with 95% confidence intervals are reported for the logistic regression. - All symptoms were analysed in separate logistic regression models: the combination of symptoms was not analysed? As it is said that the „symptom profile“ is of interest it seems that one looks at the prevalance of combination as well. Results/Tables - Table 1: Why was median, IQR included additionally for age? In the statistical analysis section it is only written that mean ± SD is shown. - It is written that for Age mean ± sd is shown but in the table it is presented as mean (sd) - What is the definition of „Non-binary“ category for Sex. This is not explained in the text. - What was the reason to categorize BMI? This is not explained in the statistical analysis section. - Table 3: reported symptoms at study entry is misleading as I think these are the symptoms they report retrospectively. These are not the symptoms they had at study entry. Additionally, the whole analysis only considers one point in time (study entry = baseline) so one can describe this in the method section and do not talk of „baseline“ anymore as this suggests that more than one visit was included. - Table 3, results of logistic regression: I think that results for symptoms with low prevalence are not robust. This can be seen for e.g. „Rash on skin“ as the CI is very wide. One can think of combining some categories if clinically meaningful. - At the end of the results section (page 7, lines 27/28) symptoms with the strongest association are listed. I think that this is only based on the Odds Ratio. However, fatigue has a similar

	OR as headache but is not listed. Instead vomiting is listed. Just looking at the OR can be misleading due to low prevalances of some symptoms and wide CIs.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1) The authors state in the abstract that their main objective is to "aid in earlier identification of SARS-CoV-2 infection" by elucidating a symptom profile for SARS-CoV-2 infection. Yet they haven't performed basic analyses to achieve their objective such as sensitivity or specificity analyses.

We thank the reviewer for this helpful comment suggesting that we include sensitivity and specificity analyses. In order to address the reviewer's comment, we have conducted these analyses and the results may be found in Supplemental Table 1. Updated text regarding these changes may be found in the Abstract (lines 82-83), Methods (lines 206-207), Results (lines 245-248), and Discussion (lines 256-259).

Line 81: Fatigue (82.9) and headache (74.9) had the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had the highest specificities among significant symptoms associated with COVID-19.

Line 206: Additionally, we included sensitivity and specificity analysis for each symptom (Supplemental Table 1).

Line 245: Fatigue (82.9), headache (74.6), and aches and pains or sore muscles (66.3) were shown to have the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had high specificity among the significant symptoms (Supplemental Table 1).
Lines 256: Discriminating symptoms for COVID-19-positivity included loss of taste and smell and bone or nerve pain as demonstrated by specificity analyses; while fatigue, headache, and aches and pains or sore muscles were shown to have the highest sensitivities among symptoms.

2) Relevance. The authors cite a few similar studies to their work but state that "However, a major limitation of studies conducted to date is the lack of comparison of patient-reported symptoms to those of uninfected individuals." The authors attention should be drawn to two Cochrane reviews which have reviewed exactly that. The most recent Cochrane review, published in Feb 2021, entitled, "Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19" included 44 studies including 26,884 participants in total. The authors should justify how this study could possible add to the vast literature that already exists on this topic. Additionally, the prior data has suggested that specific symptoms are not a very good predictors of positive test. How could this relatively small data set change that conclusion and should it?

We thank the reviewers for directing our attention to this review article. The reviewer is likely aware of the limitations to review articles, but it appears that the reviewer may be suggesting that reports of symptoms may no longer be of importance due to the completion of the cited article. If this was the reviewer's point, we respectfully disagree, for two primary reasons. First, neither the cited meta-analyses nor future pooled or meta-analyses could be completed in the absence of studies such as the one reported here. Next, we would like to draw the reviewer's attention to some of the stated

limitations of the review article she cites, including that the data included in the paper “makes it very difficult to judge the validity of the diagnostic accuracy of the signs and symptoms from these included studies”. Additionally, the authors state that additional studies “in an unselected population” are still urgently needed. Therefore, there are limitations to review articles that we are able to address herein. Finally, it is important to have representation of the unique population of Arizona in the literature regarding symptoms. Symptoms appear to vary by geographic region, SARS-CoV-2 strain, season, and other. Each paper in which these aspects differ from another is not only worthy of publication, but necessary to continue to understand the pathology of infection. We have updated the Background to include the findings from the paper cited by the reviewer (lines 120-126).

Line 120: “Additionally, a systematic review published in February 2021 aimed to assess the diagnostic accuracy of symptoms associated with COVID-19; this review identified 44 studies which in total included over 26,000 participants. The review found that among 84 symptoms, cough and fever had high sensitivities and could be used to signal for further COVID-19 testing. A limitation of this work includes possible selection bias due to their sample being selected from primarily clinical settings. Additional work examining symptoms in an unselected population is necessary to determine the syndromic presentation of COVID-19 in the general population []).

3) The text states: “Participants were asked if they had tested for the virus that causes COVID-19 with a nasal swab, throat swab, or saliva test since January 2020. Participants were classified as untested, positive or negative based on their results. Therefore, the Covid cases in this study are not lab confirmed, they are self-reported. All parts of the manuscript that refer to “lab-confirmed Covid” should be replaced with “self-reported positive Covid test.”

We respectfully disagree with this assessment. The participants in this study included two primary recruitment streams, as described in detail in the Methods (line 155-161). The primary recruitment method was via case investigations of COVID19-positive cases through a partnership with the state health department and COVID-19 testing sites and research studies. In contrast to the reviewer’s opinion, both of these methods yield laboratory confirmed results.

Line 155: “Briefly, the primary sources of recruitment have been through case investigations in a partnership with the Arizona Department of Health Services and other research studies and testing sites at the University of Arizona and Arizona State University, both of which have allowed for inclusion of laboratory-confirmed COVID-19 positive and negative participants. By October 1st, 2021, a total of 493 COVID-19-positive participants had been recruited through health department case investigations and 901 through our partnerships with studies and testing sites in Arizona.”

4) The “untested participants” should be split by those that answered “yes” to the question “Since January 1, have you experienced a sudden illness that led you to believe that you had Covid 19” and those that did not in all analyses and tables.

We have updated the analysis to reflect only those participants who have reported symptoms. Untested participants who did not experienced any symptoms have removed. The text has been updated to reflect these changes (line 66; lines 182-187).

Line 66: “… among all participants who reported symptoms within a large, prospective cohort study.”

Line 181: “All participants were first asked, “Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?” If they answered “yes”, all participants, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. Participants who respond “no” are not asked about symptomology and were not included in this analysis.”

5) Sensitivity and Specificity analyses should be performed

In order to address the reviewer's comment, we have conducted sensitivity and specificity analyses, which have been included in Supplemental Table 1.
Updated text regarding these changes may be found in the Abstract (lines 82-83), Methods (lines 205-206), Results (lines 245-248), and Discussion (lines 256-259).

Line 81: Fatigue (82.9) and headache (74.9) had the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had the highest specificities among significant symptoms associated with COVID-19.

Line 206: Additionally, we included sensitivity and specificity analysis for each symptom (Supplemental Table 1).

Line 245: Fatigue (82.9), headache (74.6), and aches and pains or sore muscles (66.3) were shown to have the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had high specificity among the significant symptoms (Supplemental Table 1).
Lines 256: Discriminating symptoms for COVID-19-positivity included loss of taste and smell and bone or nerve pain as demonstrated by specificity analyses; while fatigue, headache, and aches and pains or sore muscles were shown to have the highest sensitivities among symptoms.

6) Please report p-values for BMI and Covid symptoms analyzed as continuous variables.
These analyses have been updated in Table 2 and in the text of the Methods (line 192-194). We have updated Table 2 to include the p-values for BMI as a continuous variable.

Line 192: "From these data, we calculated body mass index as (kg/m^2), and categorized participants as having a BMI of <25, >25-29.9, and > 30, to aid in clinical interpretation, as well as reported BMI as a continuous variable (Table 2)."

7) Please add BMI and smoking to the multivariable model
We thank the reviewer for this comment and have updated the model to include BMI and smoking status. We have updated the text to reflect this change (Lines 205, 238)

Line 205: "… after adjusting for age, sex, ethnicity, BMI, and smoking status."

Line 238: "After adjusting for age, ethnicity, sex, BMI, and smoking status, COVID-19 positive participants were more likely than negative participants to experience…"

8) The text states: Participants, regardless of COVID-19 test status, were asked, "Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?" If they answered "yes", all participants, regardless of case status, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field.

"If they answered "yes" should be omitted and the Sentence should start, "All participants, regardless of case status, were asked…

We thank the reviewer for catching this error. The text has been changed in the Methods (line 187-189).

Line 187: “Regardless of symptom status, all participants were then asked if they had tested for the virus that causes COVID-19 with a nasal swab, throat swab, or saliva test since January 2020.”

9) Limitations should be expanded to include the self-reported nature of this data. Question order bias should be addressed (ie were those that answered "no" to Covid illness or to a test going to be as liking to describe symptoms), as should non-response bias.

We have included additional text in the Discussion (line 351-358) discussing the potential limitations of this design.

Line 351: “It is also important to acknowledge that participants who indicated not experiencing symptoms that led them to believe they had COVID-19, regardless of testing status, were not asked to indicate which symptoms they had experienced. These participants were not able to directly indicate that they experienced no symptoms from the provided list; however, because the majority of CoVHORT participants who undergo testing for COVID-19 enroll at a time point after receiving their test results, we believe that the likelihood that participants who indicate not experiencing symptoms actually experienced symptoms from our list offered to participants who indicate experiencing symptoms is low.”

Reviewer #2
1) If I understood it correctly, the cohort was launched at the end of May 2020 and participants were asked about the time period since January 1, so participants had to remember which symptoms they had experienced and when. One can imagine that participants with confirmed diagnosis some time ago does not remember as good as participants with a confirmed diagnosis shortly before being included in the study. Have you asked for the date of the test and was this included in the analysis? Is this the variable “days since symptoms began” (Table 2, variable not defined in the text)? Additionally, subjects could have had two tests since January 1 (e.g. two negative ones). Have you asked for that?
Thank you for this helpful comment. As the reviewer points out, our participants may have had more than one test. However, if a participant in the study was negative and then tested positive, they received the symptom questionnaire again, and only symptoms associated with the positive test were considered in this analysis. Participants who may have had more than one test are prompted to enter the test associated with the symptoms they experienced; if a participant reported more than one test, we used the symptoms associated with their positive test date. We have asked for the date of the test and date of symptom onset in our survey and have noted the date they reported their test. Using this data, we are able to calculate days between symptom onset/date of test and survey completion; this data is shown in Table 2 and defined in the Methods section of the text (Line 227 - 229). There was no significant difference found in days since test or symptom onset by symptom severity group.

Line 227: “We assessed days since symptom onset and days since test date with the survey completion date and found no significant difference between symptom groups (Table 2)”

2) It seems to me that not all participants were asked about symptoms but only those who believe to have had COVID19 as written on page 6 (lines 5/6). If this is the case, then the objective has to be restricted to this subgroup as one only knows about symptoms from that group. One has to show the frequency for answering yes to that question stratified by case status.

We thank the reviewer for this helpful comment. While we ask all participants whether they experienced symptoms that led them to believe that may have had COVID-19, we only ask the full symptom list at baseline to those who answer “yes”. In response to this as well as a comment from the other reviewer, we have changed the text cited here to clarify that only participants who responded they had symptoms that led them to believe they may have COVID19 were included in this

analysis (Line 182-187). We also added additional text (lines 351-358) discussing the potential limitations of this design. Additionally, we have updated the objective of the paper to reflect this change.

Line 182: “All participants were first asked, “Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?” If they answered “yes”, all participants, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. Participants who respond “no” are not asked about symptomology and were not included in this analysis.”

Line 351: “It is also important to acknowledge that participants who indicated not experiencing symptoms that led them to believe they had COVID-19, regardless of testing status, were not asked to indicate which symptoms they had experienced. These participants were not able to directly indicate that they experienced no symptoms from the provided list; however, because the majority of CoVHORT participants who undergo testing for COVID-19 enroll at a time point after receiving their test results, we believe that the likelihood that participants who indicate not experiencing symptoms actually experienced symptoms from our list offered to participants who indicate experiencing symptoms is low.”

3) Additionally, I was wondering if participants with no symptoms from Table 2 all belong to those answering “no “but never indicated directly to have had no symptoms. And for the untested group: I think that nearly all here answered “no”and were not asked about symptoms as one would have had performed a test if one believe to have COVID19. Please specify.

We thank the reviewer for the comment and agree with their assessment. We have revised the manuscript to restrict our sample to any participant who has reported experiencing symptoms or has tested positive. All participants in Table 2 have tested positive or have who indicated having symptoms. We feel as though it is important to highlight positive people who were asymptomatic in Table 2 to provide additional demographic information about this group. Anyone that was untested and did not experience symptoms has been removed from this analysis (line 182). The untested group in Table 3 represent all those who have experienced symptoms.

Line 182: “All participants were first asked, “Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?” If they answered “yes”, all participants, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. Participants who respond “no” are not asked about symptomology and were not included in this analysis.”

4) Restructure the sections in the abstract as the results should be described solely in the Results section and not in the Objective, Participants, Outcome Measure section. For example, the Outcome Measure section should describe the outcomes (the clear definition of the primary and secondary outcomes is also missing in the main text; this has to be added).

We thank the reviewer for this comment and have updated the abstract to reflect this suggestion (Line 65-87). We have added the primary and secondary outcomes to the main text in the Discussion section (Line 250)

Line 65: Objective: To elucidate the symptoms of laboratory-confirmed COVID-19 cases as compared to laboratory-confirmed negative individuals and to the untested general population among all participants who reported symptoms within a large, prospective cohort study. Setting and Design: This work was conducted within the framework of The Arizona CoVHORT, a longitudinal prospective cohort study conducted among Arizona residents. Participants: Eligible participants were

any individual living in Arizona and were recruited from across Arizona via COVID-19 case investigations, participation in testing studies, and a postcard mailing effort. Primary and Secondary Outcome Measures: The primary outcome measure was a comparison of the type and frequency of symptoms between COVID-19 positive cases, tested but negative individuals, and the general untested population who reported experiencing symptoms consistent with COVID-19. Results: Of the 1,335 laboratory-confirmed COVID-19 cases, 180 (13.5%) reported having no symptoms. Of those that did report symptoms, the most commonly reported were fatigue (82.2%), headache (74.6%), aches, pains, or sore muscles (66.3%), loss of taste or smell (62.8,) and cough (61.9%). In adjusted logistic regression models, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 12.1; 95% CI 9.6-15.2); bone or nerve pain (OR 3.0; 95% CI 2.2 - 4.1), headache (OR: 2.6; 95% CI 2.2-3.2), nausea (OR: 2.4; 95% CI 1.9-3.1), or diarrhea (OR: 2.1; 95% CI 1.7-2.6). Fatigue (82.9) and headache (74.9) had the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had the high specificities among significant symptoms associated with COVID-19. Conclusion: When comparing confirmed COVID-19 cases with either confirmed negative or untested participants, the pattern of symptoms that discriminates SARS-CoV-2 infection from those arising from other potential circulating pathogens may differ from general reports of symptoms among cases alone.
Line 250: “We assessed the type and frequency of symptoms between COVID-19 positive cases, tested but negative individuals, and the general untested population who reported experiencing symptoms consistent with COVID-19.”
5) Lines 18/19: one has to specify that differences in reported symptoms are identified and that this is only between positive and untested as well as positive and negative. The comparison of untested and negative is not reported.

We thank the reviewer for this comment. We have revised the Methods (line 196) to address this comment.

Line 196: “Data were analyzed to describe the COVID-19 symptoms, estimate the prevalence of individual symptoms, and identify differences among COVID-19-positive cases compared to COVID-19-negative individuals and untested participants.”

6) Lines 22/23: it has to be specified that the comparison by ordered logistic regression is only with respect to the COVID-19 positive study participants

Thank you for this helpful comment. We have revised the Methods (lines 200-202) to address this comment.

Line 200: “Among those who tested positive for COVID-19, we compared the participant characteristics at baseline and number of symptoms (0 symptoms, 1-6 symptoms, 7-9 symptoms, 10-16 symptoms) using ordered logistic regression.”

7) I do not see that nonparametric analogs were used. Please indicate here when and with respect to which analysis would you use which nonparametric analog.

We thank the reviewer for catching this error. Non-parametric analogs were not used in the final draft of the manuscript and we have removed this sentence from the methods section.

Lines 22/23: what was the reason for these symptom categories (0, 1-6, 7-9, 10-16)?

We categorized number of symptoms as a proxy to measure severity of disease. We are interested in exploring how various factors differ by severity of disease course and believe number of symptoms can estimate the severity of disease course for participants. Additionally, we have added in

participant's self-rated severity score to Table 2. This question asked participants to rank the self-perceived severity of their illness; there is a significant difference between number of symptoms reported and self-perceived severity of illness.

9) Line 26: one has to add that age, sex and ethnicity were included additionally in the logistic regression model. Please justify why those variables were considered as confounder and not e.g. BMI.

Age, sex, and ethnicity were selected as confounders in this model to reflect the existing literature on COVID-19. References from the introduction (lines 99-116) demonstrate how and why these particular variables were adjusted for in the model. We did not originally include BMI or smoking as confounders as we feel there is not yet enough evidence to link these measures to known risk factors for COVID-19. Due to reviewer feedback, we have updated the model to include BMI and smoking status. However, after recreating Table 3 to include BMI and smoking status in the model we found that estimates did not change by more than 10%, and we would prefer to restrict our model to include only age, sex, and ethnicity, if possible.

10) Indicate how missing data were handled

Missing data was removed from all analyses. All variables in the analysis, with the exception of days since symptom onset, had less than 5% missing data. Missing data for days since symptom onset is due to data entry errors from participants; these data were excluded as we were unable to verify correct dates for these participants.

11) Indicate that this is an explorative analysis, and that p-values have to be interpreted descriptively (no confirmatory value). Why reporting p-values for the ordered logistic regression and not for the logistic regression. Indicate that Odds Ratios with 95% confidence intervals are reported for the logistic regression.

We have included the following clarification in the statistical methods section (lines 200-205). We chose to use point estimates and confidence intervals to demonstrate significance of symptoms as we feel this conveys the same information as p-values in Table 3.

"Among those who tested positive for COVID-19, we compared the participant characteristics upon study entry and number of symptoms (0 symptoms, 1-6 symptoms, 7-9 symptoms, 10-16 symptoms) using ordered logistic regression and report p-values to explore factors associated with increasing severity A logistic regression model was fit for each symptom to measure the association, as measured by odds ratios and 95% confidence intervals, with COVID-19-positive status after adjusting for age, sex, ethnicity, BMI, and smoking status."

We have included the following in the Discussion (lines 367-368)

Line 367: This was an exploratory study, with a large number of statistical tests, and therefore care should be taken when considering p-values.

12) All symptoms were analyzed in separate logistic regression models: the combination of symptoms was not analyzed? As it is said that the "symptom profile" is of interest it seems that one looks at the prevalence of combination as well.

The reviewer brings up a great point. The symptoms in this manuscript were assessed individually and we cannot comment on a “symptom profile” as we have not conducted a cluster analysis. We have removed all references to a “symptom profile” (lines 324, 326) in this paper.

12) Table 1: Why was median, IQR included additionally for age? In the statistical analysis section it is only written that mean ± SD is shown.

We thank the reviewer for pointing out this inconsistency. We have removed median and IQR from Table 1.

13) It is written that for Age mean ± sd is shown but in the table, it is presented as mean (sd)

We thank the reviewer for catching this error. We have ensured that the notation is now consistent throughout the manuscript.

14) What is the definition of “Non-binary“ category for Sex. This is not explained in the text.

We thank the reviewer for pointing out the need for clarification surrounding the term non-binary. To address this comment, we have added the definition of non-binary gender in the footnotes of Tables 1 and 2; further, we have clarified that we are referring to gender and not biological sex.

15) What was the reason to categorize BMI? This is not explained in the statistical analysis section.

We thank the reviewer for this clarifying question. Body mass index was categorized using established risk categories for body size in order to improve clinical translation of the results. We have added the additional explanation in line 192.
Line 192: “From these data, we calculated body mass index as (kg/m^2), and categorized participants as having a BMI of <25, >25-29.9, and > 30, to aid in clinical interpretation, as well as reported BMI as a continuous variable (Table 2).”

16) Table 3: reported symptoms at study entry are misleading as I think these are the symptoms they report retrospectively. These are not the symptoms they had at study entry. Additionally, the whole analysis only considers one point in time (study entry = baseline) so one can describe this in the method section and do not talk of “baseline” anymore as this suggests that more than one visit was included.

The reviewer is correct in that symptoms must be reported retrospectively, after a diagnosis of COVID-19. Therefore, in response to this comment, we have omitted all usage of the word “baseline” from the manuscript
.
17) Table 3, results of logistic regression: I think that results for symptoms with low prevalence are not robust. This can be seen for e.g. “Rash on skin” as the CI is very wide. One can think of combining some categories if clinically meaningful.

We thank the reviewer for this thoughtful comment. As we prepared the revised manuscript, we considered whether collapsing some of the symptoms would aid in the interpretation of the results. However, we ultimately decided that as this is a new pathology, we wanted to provide as much detail as possible related to the reported symptoms of our study participants. However, we acknowledge the

need to be cautious in interpreting findings of rarer outcomes, and we agree with the reviewer that it will be important to assess symptom clustering in future manuscripts.

18) At the end of the results section (page 7, lines 27/28) symptoms with the strongest association are listed. I think that this is only based on the Odds Ratio. However, fatigue has a similar OR as headache but is not listed. Instead vomiting is listed. Just looking at the OR can be misleading due to low prevalence of some symptoms and wide CIs.

We thank both of the reviewers for highlighting that sensitivity and specificity analyses would strengthen the manuscript, and this recommendation most certainly has done exactly that. We have added Supplemental Table 1 to the paper, which demonstrates the sensitivity and specificity of each symptom. . Updated text regarding these changes may be found in the Abstract (lines 82-83), Methods (lines 205-206), Results (lines 245-248), and Discussion (lines 256-259).

Line 81: Fatigue (82.9) and headache (74.9) had the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had the highest specificities among significant symptoms associated with COVID-19.

Line 206: Additionally, we included sensitivity and specificity analysis for each symptom (Supplemental Table 1).

Line 245: Fatigue (82.9), headache (74.6), and aches and pains or sore muscles (66.3) were shown to have the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had high specificity among the significant symptoms (Supplemental Table 1).
Lines 256: Discriminating symptoms for COVID-19-positivity included loss of taste and smell and bone or nerve pain as demonstrated by specificity analyses; while fatigue, headache, and aches and pains or sore muscles were shown to have the highest sensitivities among symptoms.

VERSION 2 – REVIEW

REVIEWER	Kirchner, Marietta UniversitätsKlinikum Heidelberg, Medical Biometry
REVIEW RETURNED	15-Nov-2021

GENERAL COMMENTS	My comments have all been adressed. I have just one follow-up recommendation with respect to Table 3 and the handling of missing data: Table 3: you wrote that you have run the model with BMI and smoking status and as you found similar results would like to stick to the previous one with only age, sex, and ethnicity included as confounder. However, in the revised manuscript you show the results adjusted for all five confounder and this is also written in the "Statistical analysis Section". So I am not sure what do you want to show in the final version. Aditionally, I propose to add that the odds ratios resulted from a logistic regression model which was adjusted for the confounder in the table caption. One can add the reason for deciding on the included confounder variables in the Statistiscal analysis Section. Missing data: thank you for clarifying that missing data was not imputed. I propose to add this information to the Statistical Analysis Section and to indicate if the analysis was then based on e.g. a pairwise deletion. So the results in Table 3 are based on all subjects with complete information on the confounder variables

	which is then different if adjusting for the three or all five confounders? Sample size for the logistic regression should be added here to clarify this. I think the first three columns of Table 3 include all subjects and only the reporting of OR (last two columns) is based on the reduced sample excluding subjects with missing value or what do you mean with "missing data was removed from all analyses" (what is summarized by "all analyses" here)? Please clarify.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Dr. Marietta Kirchner, Universitäts Klinikum Heidelberg
Comments to the Author:

My comments have all been addressed. I have just one follow-up recommendation with respect to Table 3 and the handling of missing data:

Table 3: you wrote that you have run the model with BMI and smoking status and as you found similar results would like to stick to the previous one with only age, sex, and ethnicity included as confounder. However, in the revised manuscript you show the results adjusted for all five confounder and this is also written in the "Statistical analysis Section". So, I am not sure what do you want to show in the final version. Additionally, I propose to add that the odds ratios resulted from a logistic regression model which was adjusted for the confounder in the table caption. One can add the reason for deciding on the included confounder variables in the Statistical analysis Section.

We thank the reviewer for this comment. We have decided to leave the manuscript as it currently is with the confounders age, sex, ethnicity, BMI, and smoking status and have added in reasoning in the statistical analysis section. We have also updated the Table caption to reflect the reviewer's suggestion.

Line 213: "Confounders were selected based on background knowledge."

Line 415: "Table 3. Symptom characteristics and odds ratios of CoVHORT participants using a logistic regression model adjusted for case status, age, sex, ethnicity, BMI, and smoking status."

Missing data: thank you for clarifying that missing data was not imputed. I propose to add this information to the Statistical Analysis Section and to indicate if the analysis was then based on e.g. a pairwise deletion. So, the results in Table 3 are based on all subjects with complete information on the confounder variables which is then different if adjusting for the three or all five confounders? Sample size for the logistic regression should be added here to clarify this. I think the first three columns of Table 3 include all subjects and only the reporting of OR (last two columns) is based on the reduced sample excluding subjects with missing value or what do you mean with "missing data was removed from all analyses" (what is summarized by "all analyses" here)? Please clarify.

We thank the review for this comment. We have added the following sentence to the statistical analysis section for clarification.

Line 213-215: “Logistic models were performed using participants with complete data (n=1,370) for all variables in the model. Additionally, we included sensitivity and specificity estimates for each individual symptom”

Reviewer: 2
Competing interests of Reviewer: No competing interests.